Thompson Declaration

Redacted Version of Document Sought to be Sealed

Mark C. Mao, CA Bar No. 236165
Beko Reblitz-Richardson, CA Bar No. 238027
Erika Nyborg-Burch, CA Bar No. 342125
BOIES SCHILLER FLEXNER LLP
44 Montgomery St., 41st Floor
San Francisco, CA 94104
Tel.: (415) 293-6800
mmao@bsfllp.com
brichardson@bsfllp.com
enyborg-burch@bsfllp.com

James Lee (admitted *pro hac vice*)
Rossana Baeza (admitted *pro hac vice*)
BOIES SCHILLER FLEXNER LLP
100 SE 2nd St., 28th Floor
Miami, FL 33131
Tel.: (305) 539-8400
jlee@bsfllp.com
rbaeza@bsfllp.com

Amanda K. Bonn, CA Bar No. 270891
SUSMAN GODFREY L.L.P
1900 Avenue of the Stars, Suite 1400
Los Angeles, CA 90067
Tel: (310) 789-3100
Fax: (310) 789-3150
abonn@susmangodfrey.com

Attorneys for Plaintiffs

William Christopher Carmody
(admitted *pro hac vice*)
Shawn J. Rabin (admitted *pro hac vice*)
Steven M. Shepard (admitted *pro hac vice*)
Alexander Frawley (admitted *pro hac vice*)
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas,
32nd Floor
New York, NY 10019
Tel.: (212) 336-8330
bcarmody@susmangodfrey.com
srabin@susmangodfrey.com
sshepard@susmangodfrey.com
afrawley@susmangodfrey.com

John A. Yanchunis (admitted *pro hac vice*)
Ryan J. McGee (admitted *pro hac vice*)
MORGAN & MORGAN
201 N. Franklin Street, 7th Floor
Tampa, FL 33602
Tel.: (813) 223-5505
jyanchunis@forthepeople.com
mram@forthepeople.com
rmcgee@forthepeople.com

Michael F. Ram, CA Bar No. 104805
MORGAN & MORGAN
711 Van Ness Ave, Suite 500
San Francisco, CA 94102
Tel: (415) 358-6913
mram@forthepeople.com

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

CHASOM BROWN, WILLIAM BYATT, JEREMY DAVIS, CHRISTOPHER CASTILLO, and MONIQUE TRUJILLO individually and on behalf of all similarly situated,

Plaintiffs,

vs.

GOOGLE LLC,

Defendant.

Case No.: 4:20-cv-03664-YGR-SVK

DECLARATION OF CHRISTOPHER THOMPSON IN SUPPORT OF PLAINTIFFS' REQUEST FOR AN ORDER TO SHOW CAUSE

The Honorable Susan van Keulen
Courtroom 6 - 4th Floor
Date: April 21, 2022
Time: 10:00 a.m.

DECLARATION OF CHRISTOPHER THOMPSON

I, Christopher Thompson, declare:

1. Counsel for the *Brown* Plaintiffs retained me to provide technical analysis and testimony in connection with the upcoming evidentiary hearing on Plaintiffs' Request for an Order to Show Cause, including in response to the technical assertions made by Google in its opposition filing and by various Google declarants who filed statements in support of Google's opposition filing.

2. All of the statements in this declaration are true based on my analysis and personal knowledge, and I am available and if the Court permits it willing to testify on these matters during the upcoming evidentiary hearing.

3. A copy of my CV is attached as Exhibit A. As reflected in my CV, I majored in Computer Engineering and have many years of experience in computing technology. I am being compensated at a rate of $275 per hour for my work in connection with this matter, and none of my compensation is contingent on the outcome of this litigation.

4. In the course of my previous work writing software and building software systems, I have used Protocol Buffers, defined "proto" schema files, and built systems that write to the data structures defined by proto files.

5. I have reviewed each and every submission Google and the Special Master made available as part of the Special Master process, including the Plaintiffs' data and test data produced by Google, and the transcripts of the hearings before the Special Master. In addition, all documents Google produced and deposition transcripts for witnesses in this case have been made available to me pursuant to the Protective Order issued in this case.

6. I was also present at a live test demo with Google engineers and Special Master Douglas Brush on March 4, 2022. At that session, we had tested a small set of Biscotti IDs against [redacted] of the [redacted] logs.

Google's Ability To Detect Event-Level Incognito Traffic Within Its Logs

7. I reviewed Google's Opposition to Plaintiffs' Request For an Order For Google to Show Cause For Why It Should Not Be Sanctioned for Discovery Misconduct ("Google's Opposition" to "Plaintiffs' Request"), and I understand that Google is arguing that event-level Incognito usage cannot be identified.

8. Based on my analysis of the data produced by Google in this litigation, including in connection with the Special Master process, that assertion is incorrect. The data produced by Google confirms that Google can (and in fact does) detect event-level Incognito traffic within its logs.

9. I provide two simple experiments we used to demonstrate this event-level detection. This assessment is based on data produced by Google, and I worked with Plaintiffs' consultant Dr. Lillian Dai to prepare these examples.

10. In one example, we used IP addresses and user agent strings to identify event-level Incognito traffic. Because this was data produced in connection with the Special Master, the data we were able to test was limited to the named Plaintiffs and certain test accounts created in connection with that process (referred to below as our "consulting team" accounts). First, we located the user's IP address and user agent string, either from the device, GAIA account information, or from GAIA logs. From the First Iterative Search with the Special Master, a GAIA log search against [redacted], containing a [redacted] field in row 2 equal to "2454128719," which is converted[1] to IP address 146.71.8.79.[2] The same row contains user agent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/96.0.4664.55 Safari/537.36,gzip(gfe)".

[1] IP addresses may be converted using https://www.browserling.com/tools/dec-to-ip.

[2] This file lacks a Bates stamp, and is instead produced as [redacted] [redacted].csv", produced by Google to the Special Master on February 23, 2022, as part of a production named "20220223 Brown v. Google - [redacted]". Plaintiffs will be prepared to present this search result produced by Google at the evidentiary hearing.

11. Next, we located the user's Biscotti ID using this IP address and user agent string pair. In the same example, GOOG-BRWN-00826529[3] ([redacted]), GOOG-BRWN-00826530 ([redacted]), GOOG-BRWN-00826531 ([redacted]), GOOG-BRWN-00826532 ([redacted]), GOOG-BRWN-00826534 ([redacted]), GOOG-BRWN-00826535 ([redacted]), GOOG-BRWN-00826536 ([redacted]), GOOG-BRWN-00826537 ([redacted]) and GOOG-BRWN-00840745 ([redacted]) contained our consulting team's Incognito signed-out experimental data associated with Biscotti ID "2501521082151731303". GOOG-BRWN-00826130 ([redacted]) contained our consulting team's Incognito signed-out experimental data associated with Zwieback ID "0xa30eae52e9dcb304". All of these Incognito signed-out Display and Search ad logs contain the same IP address and user agent as that in the GAIA log: RemoteHost: "146.71.8.79" or client_ips: "2454128719" and UserAgent: "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/96.0.4664.55 Safari/537.36,gzip(gfe)". The user agent corresponds to a Mac device and Chrome browser.

12. We also checked the Google ad logs containing the Biscotti IDs to verify Incognito usage. Here, we checked the above-referenced [redacted], [redacted] and [redacted], as well as logs containing [redacted] associated with Biscotti ID 2501521082151731303, to verify Incognito usage. The first [redacted] logs contain the X-Client-Data Header field, and the results had no X-Client-Data Header value for the correct IP address and user agent string. And the last set of logs contain the [redacted] bit, which Plaintiffs tested in a live demo on March 4, 2022, with Google engineers in which I was present. Of the [redacted] logs tested, the result had [redacted] set as "true" for the Biscotti ID 2501521082151731303.

[3] For all Bates stamped logs referenced for this first example, these were all natively produced spreadsheets provided by Google. Plaintiffs can provide them to the Court or Google readily upon request. Regardless, Plaintiffs will be prepared to present them at the evidentiary hearing.

13. As a second example, we used Google Analytics User IDs to identify event-level Incognito traffic. First, we located the user's Google Analytics User ID ("UID"). This time from the Second Iterative Search, production "2022-03-25 Brown v. Google – Analytics [redacted] data – AEO", file [redacted] [redacted] [redacted] [redacted][4] row 2246 corresponded to Plaintiff Mr. Jeremy Davis' UID "D6E68756C7085109E0530100007F4E1E" from washingtonpost.com. Column M of the same row contained a request URL containing his CID from washingtonpost.com: [redacted].

14. Next, we located the user's Biscotti ID using the CID. In the same example, CID [redacted] was found in the file [redacted] [redacted], at row 5 and column M. The same row, in column A showed Mr. Davis' Biscotti cookie "AHWqTUkuQpT6kkO-Dw_-ua3QraXieMCgN4y9rGORTwXNcUaWhg5Y47ntF2PavJTgdkg". The embedded Biscotti ID in this cookie is shown in "2022-03-02 Brown v. Google - Decode IDE.pdf", at page 4, item 33, as [redacted].

15. We then checked Google's ad logs containing the Biscotti IDs to verify Incognito usage. The logs referenced above ([redacted], [redacted] and [redacted]) were checked to verify Incognito usage against Biscotti ID [redacted]. The results had no X-Client-Data Header value.

16. As the two above examples show, from only the First and Second Iterative Searches, the [redacted] bit developed and implemented by Google, like the X-Client-Data Header field, sit in logs that contain identifiers such as a Biscotti ID, or can be located using other identifiers such as an IP address and user agent string pair. As I discuss below in the next section, the [redacted] and [redacted] bits operate similarly,

[4] For all xlsx spreadsheets referenced for this second example, these were also all natively produced spreadsheets provided by Google. These sources, as named by Google, were identified correctly. Plaintiffs can provide them to the Court or Google readily upon request. Regardless, Plaintiffs will be prepared to present them at the evidentiary hearing.

and are contained within the GWSLogEntryProto (sometimes referred to by Google in filings as "GWS Proto") bit schema, to which GWS logs would have ready and easy access.

17. Because Google has not provided data for all three of these Incognito-detection bits (let alone any other Incognito-detection bits that might exist), or all of the associated logs and sources with their full schema, it is still unclear the extent to which different Google logs and sources can be used for similar identification purposes. Still, the above two examples illustrate at least some ways by which event-level Incognito detection and identification can be done (depending on the data retained by Google).

Google's Ability To Preserve Event-Level Incognito Traffic

18. Google's [redacted] and [redacted] bits are especially useful for identification and preservation purposes because they are GWSLogEntryProto bits. Because they exist in GWSLogEntryProto, this means that the two Google bits can be used in connection with a number of different logs. Given that these fields were built in 2017 and 2018, they could have been made "live" at the beginning of the case or "added" to any GWS log. Put differently, any process within Google that writes to a log using the GWSLogEntryProto data structure can simply write to that specific field. Writing to an existing field within a Protocol Buffer data structure is a one-line code addition. For example, Google could have easily added these Incognito-detection bits into any of the GWS logs referenced in the two examples discussed above.

19. Google's own public documentation on the Protocol Buffers library and specification explains how easy it is to write to an existing field.[5] The example from the documentation involves writing an ID to "person" message, and in C++ it is as simple as person->set_id(id); where "id" is the desired value.

20. Contrary to what Google's Opposition suggests, these Incognito-detection bits do not appear to be just for "Search logs." During the Special Master process, the [redacted] bit appeared in the schema for the [redacted] and

[5] https://developers.google.com/protocol-buffers/docs/cpptutorial#writing-a-message

[redacted], as part of the March 11, 2022 production, in a file named "2022-03-10 Brown v. Google – Fields for [redacted] Logs – AEO.xlsx."[6] This means that the logs either were already collecting data for these fields, or can simply be authorized to collect data for these fields. These are not Search-only Incognito-detection bits. Importantly, while these [redacted] Incognito-detection bits also use the X-Client-Data Header or some logic relying on the same, the bits are much smaller to store than the X-Client-Data Header field. The bits are "Boolean" in that they simply store a "yes (1)" or "no (0)" value, and would have added minimal weight to any existing log if turned on or added.

<u>Google's Incognito Traffic Detection Is At The Event-Level</u>

21. I understand that Google is arguing that the three Incognito-detection bits were built only for aggregated traffic analysis and not for event-level analysis. While that may be how Google allegedly intended to use the bits, the starting point is an event-level categorization. And the same bits can certainly be used to identify event-level data, as the above two examples from the First and Second Iterative Searches already show.

22. Perhaps more importantly, aggregated analysis still depends on event-level detection. This is an aggregation of event-level logs. That is exactly why these bits are in event-level logs. The logs using these bits that Google identified contain event-level data. While Google may use those logs to create aggregated analysis, that does not change the fact that aggregation starts with event-level Incognito-usage data. To the extent logs had been or are preserved, the logs can be used to identify Incognito-usage at an event level.

<u>Accuracy of Google's Detection of Event-Level Incognito Traffic</u>

23. I also understand that Google is asserting that these Incognito-detection methods are not necessarily "accurate." Based on my own analysis, looking at the data produced in connection with the Special Master process, that seems incorrect. The records I have seen indicate that Google is accurately detecting incognito-traffic and using these bits to identify the traffic as

[6] For xlsx spreadsheets referenced for this section, these were also all natively produced spreadsheets provided by Google. These sources, as named by Google, were identified correctly. Plaintiffs can provide them to the Court or Google readily upon request. Regardless, Plaintiffs will be prepared to present them at the evidentiary hearing.

such. To the extent there are specific instances where non-incognito traffic has been labeled as incognito traffic with these bits, that is something that could be the subject of further expert analysis, had Google preserved or produced such data.

Linkability/Joinability and Identification of Class Members

24. I also understand that Google is asserting that Incognito data is not linkable to specific users, and that Google's data cannot be used to identify class members.

25. Based on my analysis of the data produced in connection with the Special Master process, these assertions are also incorrect. As demonstrated above, the data produced by Google can be linked to specific users, who can be identified as class members.

26. These are issues where additional data, had it been preserved by Google, would have provided additional proof on these points, allowing for the identification of additional users of Chrome Incognito mode during the alleged class period.

27. The linkability of these records is also something that Google's own employees recognized during the class period. [redacted] *See* McClelland Ex. 15, GOOG-CABR-05256755 at -759; McClelland Tr. at 212:13-212:24. It is also possible for Google to join separate zwieback cookies between different incognito sessions [redacted] McClelland Tr. at 209:11-209:24; McClelland Ex. 17, GOOG-CABR-00799341. A true and correct copy of the relevant excerpts and exhibits from the deposition is attached hereto as Exhibit B.

28. I have also reviewed Google documents stating that Google logs an encrypted signed out identifier in its personal logs, and retains the encryption key for [redacted] days. GOOG-CABR-04773853, -54, -67, -88. Google employees understood that retaining the encryption key provides a mechanism for Google to link signed-in activity associated with a Google account to signed-out activity logged with the signed-out identifier. GOOG-CABR-03652549, -552-53. True and correct copies of the relevant excerpts from these two documents are attached hereto as

Exhibits C and D respectively.[7]

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct. Executed this 11th day of April, 2022, at Nolensville, Tennessee.

/s/ Christopher Thompson

[7] These lengthy documents produced by Google are cited correctly, and only excerpts are attached hereto. Plaintiffs can provide these documents to the Court or Google readily upon request. Regardless, Plaintiffs will be prepared to present them at the evidentiary hearing.